# EXHIBIT D

**IN THE UNITED STATES DISTRICT COURT**
**FOR THE DISTRICT OF MARYLAND**

| | | |
|---|---|---|
| IN RE MICROSOFT CORPORATION ANTITRUST LITIGATION | ) | |
| This Document Relates to: | ) | MDL Docket No. 1332 |
| *Burst.com, Inc. v. Microsoft Corp.* | ) | Hon. J. Frederick Motz |
| Civil Action No. 1:02cv02952 | ) | |

**DECLARATION OF BRIAN VON HERZEN**

1. My name is Brian Von Herzen. I reside at 675 Fairview Dr., #246, Carson City, Nevada 89701. I am currently employed by Rapid Prototypes, Inc. I have been retained as an expert by Microsoft Corporation in connection with this litigation.

2. I have completed two expert reports as part of my work in this case. The first is titled "EXPERT REPORT OF BRIAN VON HERZEN REGARDING U.S. PATENT NO. 5,995,705, NO. 5,164,839, AND NO. 4,963,995" ("Expert Report") and is attached as Exhibit 1. My second expert report in this case is titled "OPPOSITION EXPERT REPORT OF BRIAN VON HERZEN RELATING TO CLAIM CONSTRUCTION"(Opposition Report") and is attached as Exhibit 2.

3. In my Expert Report, I provided an explanation of some of the technology involved in this case, which I incorporate by reference into this declaration.

4. Both sound and light as they exist in the world are in analog form and people perceive them in this analog form. What people perceive as sound is actually fluctuation in air pressure. What people perceive as light is actually certain electromagnetic radiation. Electronic devices can capture and reproduce these analog signals. For example, a microphone captures analog audio and a video camera captures analog light. Both then produce an analog electric signal that represents the original audio or video.

5. The analog representation of audio or video is a continuous wave that at any point in time represents, for audio, the magnitude of the air pressure and, for video, the magnitude of the light. Analog representations are hard to process and hard to store and to transmit accurately. Because the analog waveform itself represents the audio or video, it is hard to separate noise from the audio or video signal.

6. To solve these problems, audio and video signals are often represented as digital signals. Transforming an audio or video signal into digital form requires sampling the analog wave at discrete times and discrete levels and recording the level of the wave at those times – the more samples and levels the better and more accurate the digital representation of the original analog signal. However, using more samples and levels increases the amount of digital data required to represent the original audio or video.

7. The longer the audio or video program, the more digital data is required to represent it. CD quality stereo audio, for example, uses about 44,000 samples per second with each sample being at least 16 bits for the left side and 16 bits for the right side. Thus, an hour of CD quality stereo requires at least 5.0 gigabits.

8. Digital video comprises a series of still pictures, called frames. Video appears to show motion because each frame is slightly different than the one before and after it. When the frames are displayed one after another (as described in the patents, generally at 30 frames per second), the viewer experiences a sense of uninterrupted motion, called "motion fusion."

9. In standard television quality video, each frame has as many as 480 horizontal lines of pixels and approximately 720 pixels per line. Thus, a single frame of standard definition TV (SDTV) has as many as 345,600 pixels. SDTV in the United States uses approximately 30 frames per second of video, meaning over 10 million pixels per second. At 24 bits per pixel, SDTV may include nearly 9 million bits per frame, and nearly 250 million bits per second.

10. Because of the huge data volumes needed to store audio or video content, transmitting these files across a network can take a very long time. Moreover, because file size is partially dependent on the length of the content, longer content takes more time than shorter content to transmit. Thus, first transmitting an entire audio or video file and then playing the content was impractical for content longer than a few seconds because doing so required a user to wait a long time before the content started to play.

11. Streaming allows audio or video content to be played as the file is received, eliminating the need to receive all the entire content before starting to play it. Conceptually, streaming is like conventional television. One watches a television program as it is received. With streaming, the content begins to play after the system receives only some minimum amount of data. Thus, the length of the content no longer affects the amount of time before the user starts to enjoy the content.

12. All data stored in digital electronic form is represented by a series of bits where each bit is either a "1" or a "0." As explained above, a very large number of ones and zeros are necessary to represent an audio or video signal.

13. Data compression is the process of reducing the amount data necessary to represent some information such as a program, database data, audio, or video. Compression is useful to reduce the space needed to store data and the transmission time needed to send it. For a given transmission rate, compressed data usually requires less time to transmit than the original uncompressed data because it is smaller.

14. In 1988 (and today), there are many kinds of data compression. Lossless data compression compresses data so that the exact original data can be recovered. Lossless data compression is necessary for data, like programs or databases, where the original data must be recoverable. Lossless compression, however, can compress data only so small. In many cases, such limited compression is insufficient.

15. When exact recovery of the original data is not necessary, lossy compression can provide greater compression than lossless compression. Lossy compression compresses data by discarding some of the original data, leaving only an approximation of the original data recoverable from the compressed data. Lossy compression is very useful for audio and video content. By accounting for the way humans hear and interpret visual information, lossy compression can selectively discard audio and video data, leaving decompressed audio or video data that is perceived to be similar to the original and minimizing the effect of not being able to recover all the original data.

16. Because it discards some of the original data, lossy compression allows control over the amount of compression and thus the size of the compressed data. In effect, the lossy compression algorithm can be instructed how much data to throw away and thus what size to make the compressed data.

17. This feature of lossy compression facilitates time compression of data. Time-compressing data involves compressing data, usually time-variant data like audio or video, into a particular size. The size of the time-compressed data is chosen based on the intended use of the data. For example, if the time-compressed data must be sent over a transmission link at 56 kilobits/second (a common modem speed), it must be compressed small enough to fit into that time slot.

18. Video data is often compressed using a combination of two techniques, each of which can be either lossy or lossless, but are generally lossy. Intra-frame compression compresses each frame of the video, thus reducing its size and the time it takes to transmit. To obtain additional compression, inter-frame compression reduces the size of the video data by noting the differences between each frame, which are generally small, and storing or transmitting only the variations from one frame to the next.

19. After audio and video data is compressed, it must be assembled into a known format. Common audio formats include wav, mp3, and Windows Media Audio (wma). Common video formats include MPEG2 (used for DVDs, digital cable television, and satellite systems) and Windows Media Video (wmv).

20. The patents at issue do not describe either a packet-switched network or the Internet. The Internet is a packet-switched network. These patents also do not describe using

packets. The patents do not address the technical obstacles necessary to use a packet-switched network such as addresses, headers, out of order packets, latency, and lost packets. The system described in the patents would need to be redesigned to work with a packet-switched network, such as the Internet. The system described in the patents does not need to address these technical obstacles because it is designed to work over a direct connection, such as a circuit-switched network.

21. Both standard telephone lines and fiber optic telephone lines provided by the telephone company provide essentially a wired connection between two points. The telephone company provides a constant connection over which only a certain amount of data can be transmitted per second, *i.e.*, a fixed "bandwidth." When one orders a telephone line, whether standard or fiber optic, one receives from the telephone company a guarantee of a certain level of service at a requested (and paid for) bandwidth. This connection provided by the telephone company is commonly referred to as "circuit-switched" because the telephone company creates a circuit between the two parties using switches internal to the telephone network. Similarly, a "point-to-point" link represents a circuit between two parties that does not require a switch.

22. In 1988, local and long distance telephone companies could provide fiber optic telephone lines; when a telephone customer used a fiber optic connection, it was typically a leased line. There was no such thing as a dialed fiber optic telephone line. Leased connections provide a continuous connection between two locations set up when the line is ordered. A leased connection acts as a wire between two parties that operates at the bandwidth purchased from the telephone company.

23. The patents at issue do not describe using a circuit-switched connection as part of a packet-switched network. The patents also do not describe using a packet-switched network to implement a circuit-switched connection.

24. The patents at issue describe processing a certain type of video signal. The patent describes this signal as providing 90,000 pixels (300 x 300) per frame with 21 bits per pixel providing the color information. According to the patents, these frames are received at the rate of 30 frames/second. Col. 4, Lines 57-63. This frame rate is the same as standard television, although the number of pixels per frame is less than standard television.

25. The patents at issue also describe using a small RAM 29 to hold at least two full uncompressed frames of data at a time for processing. Col. 5, Lines 46-50. The disclosed system lacks the required space in RAM 29 to store the uncompressed video frames for a program if it cannot process 30 frames/second while receiving video at that rate.

26. In 1988, software running on a general-purpose microprocessor or general-purpose computer could not have compressed a video signal of 300x300 pixels and 21 bits per pixel at the frame rate described in the patent of 30 frames/second. As explained in an IEEE article from September 1988 (Luther, "You are there ... and in control," IEEE Spectrum (September 1988), even a super-computer with 64 processors (a typical personal computer contains a single processor) required 3 seconds to compress just one frame of data. The system in the patents requires compressing 30 frames/second. The same article goes on to say that a general-purpose computer of the time such as a DEC VAX computer would take 30 seconds to compress each frame of the video.

27. Although there was no realistic way to have software in 1988 compress video data at the 30-frames/second rate described in the patents, a hardware integrated circuit could compress video data at that rate. For example, Intel's DVI hardware technology, publicly demonstrated in 1987, could compress video at 30 frames/second.

28. Choosing a hardware or software implementation of an algorithm depends on many factors, including the need for adequate performance. Often the two are not interchangeable when real-time computation is required. In 1988, hardware and software were less interchangeable for real-time applications because the software to hardware performance ratio was even more extreme than it is today. But even today, implementing a compression algorithm in hardware or software involves different strategies and dramatically different performance levels.

29. Software may offer advantages over hardware in certain instances, albeit at reduced performance. As explained above, software compression could not be performed in real time – *i.e.*, as video was being received – in 1988. But software implementations can be more flexible because their programming can be changed. Software also performs some tasks differently than hardware, influencing how one would implement a compression algorithm. Software can process conditional statements efficiently, allowing implementations to handle many alternatives intelligently and only process data relating to the most promising alternatives. Software can handle special cases and exceptions using such conditional execution. Hardware must implement each special case to be handled using dedicated hardware resources. These resources increase the area and cost of the integrated circuit, driving hardware implementations to use fewer conditionals than might be optimal in a software implementation.

30. Hardware is generally much faster than software. Hardware implementations, however, are usually less flexible. Implementing an algorithm in hardware requires different optimization strategies than doing so in software. Pipelined hardware implementations do not efficiently allow for conditional branches, which can be straightforward to implement in software. But hardware, because of its speed, allows a brute-force approach where all possible results are tested for the best outcome. Hardware allows for the massive replication of computational structures in a large array. Different circuits within such a large replicated array of hardware elements work concurrently on different parts of a video frame. These brute-force approaches use simple algorithms, massively replicated, to achieve performance levels required for real-time applications, especially in the 1988 time frame.

31. The '995 patent describes using several types of long-distance communication links, a fiber optic telephone line and a telephone line. The '705 patent also describes using a microwave link. A microwave link is a form of radio communication. In 1988, a fiber optic telephone line was a leased line that connected two locations over a fixed-bandwidth connection through the telephone network. The telephone network in 1988 was circuit switched and used time-division multiplexing to allow multiple connections to share trunk lines.

32. I understand "time compressed" in 1988 to mean an audio/video program compressed sufficiently so that the transmission time is reduced to a predetermined time. "Time compress" is a well-known term that means to compress an audio/video program so that the transmission time is reduced to a predetermined time. Time compression is commonly used with shared communications links where each user receives only a fraction of the link's total transmission capacity. The predetermined time is important because audio/video programming

is often allocated a percentage of time slots when transmitted over a circuit-switched network on a time-division multiplexed channel. If the audio/video doesn't fit within its predetermined fractional bandwidth, then it cannot be transmitted successfully over the multiplexed channel. Time compression is also used over fixed bandwidth transmission links to ensure that the data can be transmitted over the link within the necessary time. Time compression does not work over the Internet.

33. Walter (US Patent 4,506,387) discusses transmission of video over fiber optic cables. Walter also discusses transmission of video material faster than real-time for later viewing:

> "By utilizing inter-frame differential pulse code modulation, each second of video program playing time yields about 44 megabits. Further, according to the present state of the art, 650 megabits per second can be transmitted on a single wavelength, and since in the present embodiment there are 16 optical data channels in the four fiber optic lines 56, 58, 60, 62, the total transmission rate is 10,400 megabits per second. Therefore, a two hour movie can be transmitted in about 31 seconds (7,200 seconds.times.44 megabits per second /10,400 megabits per second)."

(Walter, US patent 4,506,387, Col. 7, Lines 37-44) Walter describes a broadcast cable system with video programming on demand. The video programming is stored and can be broadcast in a non-real-time basis to a user. The source video programming is compressed using inter-frame differential pulse code modulation. Inter-frame differential pulse code modulation (IFDPCM) is a form of data compression. In a nutshell, IFDPCM takes the differences between frames and encodes the difference at each pixel as a binary number. Walter does not associate any time with the compressed data other than the play time. The resulting compressed video is stored in memory modules. Walter describes transmitting the compressed video to users over optical fiber. The transmission to the user in Walter's patent occurs in less than the video's play time.

Walter describes sending a two-hour video in 31 seconds. Such video source information is compressed in digital form only a fraction of the size of the uncompressed original. Walter uses a conventional data compression technique and not time compression. Walter's compressed data differs from time-compressed data in that it compresses data but does not specify a required size of the compression results. In contrast, time compression specifies a required maximum size of the compression results.

34. The claim language for the "input means for receiving audio/video source information" in the '995 and '705 patents does not identify either structure for performing the claimed function or a known class of structures for performing the claimed function. The structure disclosed in the patents for performing this function includes the video line or camera input line 15, aux digital input port 17, fiber optic input/output port 18, or TV RF tuner input port 16 combined with selector switches 35, 36, and 37. The selector switches described in the patents are mechanical switches. No other structure described in the patents performs the claimed function.

35. The claim language for the "input means for receiving audio/video source information as a time compressed representation thereof" in the '995 patent does not identify either a structure for performing the claimed function or a known class of structures for performing the claimed function. The structure disclosed in the patents for performing this function is a fiber optic port and a mechanical selector switch and the auxiliary digital input port. No other structure described in the patents receives compressed data.

36. The claim language for the "compression means" in the '995 and '705 patents does not identify either structure for performing the claimed function or a known class of

structures for performing the claimed function. In the '995 patent, the structure disclosed as performing the compression means' function is an AMD 7971 hardware compression chip. The '995 patent mentions two algorithms for compression video data, the CCITT Group IV facsimile compression and inter-frame compression. The patent discloses structure for performing the Group IV compression, the AMD 7971, but does not disclose any structure for performing inter-frame compression. No other structure described in the patents performs the claimed function.

37. In 1988, software would not have been able to perform inter-frame compression at the 30-frames/second rate described in the patents. The patents do not describe using software to perform compression.

38. The '995 and '705 patents mention two microprocessors, CPU 28 and CPU 31. The patents do not describe either of these microprocessors as performing any compression. The patents do not describe any software compression. The CPU 28 controls, in addition to other elements, the AMD 7971 compression chip. The specification of the '995 patent describes the CPU 28 as working with the AMD 7971 compression chip but never describes the CPU 28 as compressing any data. Col. 5, Lines 46-49 ("The CPU (Central Processing Unit) 28 is a microprocessor which controls the digitization process of VCU 12. CPU 28 works with controller 27 to control and communicate with the other elements of the VCU.")

39. The '705 patent does not identify any structure for performing the function of the compression means. Like the '995 patent, the '705 patent mentions Group IV and inter-frame compression. The '705 patent does not describe any structure as performing either compression algorithm. In the '705 patent, no particular structure is identified that performs the function of compression means.

40. Many electronic devices contain microprocessors and RAM. All these devices are not computers. For example, VCRs generally have both a microprocessor and RAM, but a VCR is not a computer. The other electronic items also contain microprocessors and RAM but are not computers, including microwave ovens, calculators, ovens, washers, and dryers.

41. The '705 patent describes a compressor/decompressor block 26. Block 26 does not disclose any structure. This block is shown on Figure 2, which appears distinctly as a hardware block diagram, as evidenced by the other hardware blocks, including RAM, CPU, D/A and A/D. Therefore, I have to conclude that Block 26 is some sort of hardware chip. Additionally, block 26 is not a programmable microprocessor. Nowhere in the specification is block 26 described as a programmable microprocessor. The figure is inconsistent with a conclusion that block 26 is software. To communicate that block 26 was a microprocessor to professionals in this field, one would have so indicated and would not have made separate boxes for 26 and microprocessor 27.

42. The claim language for the "random access storage means" in the '995 and '705 patents does not identify either structure for performing the claimed function or a known class of structures for performing the claimed function. The structures disclosed in the patents for performing this function are DRAM and SRAM semiconductor memories and optical disc memories. No other structure described in the patents receives compressed data.

43. The claim language for the "storage means" in the '705 patent does not identify either structure for performing the claimed function or a known class of structures for performing the claimed function. The structure disclosed in the patents for performing this function is a DRAM and SRAM semiconductor memories, optical disc memories, bubble

memories, digital paper, and magnetic disks. No other structure described in the patents performs the claimed function.

44. The claim language for the "output means" in the '995 patent and the "transmission means" in the '705 patent does not identify either structure for performing the claimed function or a known class of structures for performing the claimed function. The structure disclosed in the '995 patent for performing the claimed function is a fiber optic port connected to a fiber optic telephone line and a mechanical selector switch. In the '705 patent, the specification also includes a microwave transceiver. No other structure described in the patents performs the claimed function. The telephone line and modem discussed in the patents do not correspond to this claim element because the patents do not link these structures to the faster-than-real-time transmission required of the "output means" and "transmission means."

45. The claim language for the "decompression means" in the patents does not identify either structure for performing the claimed function or a known class of structures for performing the claimed function. The structure disclosed in the patents for performing the claimed function is the same as the structure corresponding to the compression means. In the '995 patent, this structure is the AMD 7971 hardware compression chip, which performs compression and decompression, and in the '705 patent, no structure is disclosed as performing the function of the decompression means. No other structure described in the patents performs the claimed function.

46. The claim language for the "editing means" does not identify either structure for performing the claimed function or a known class of structures for performing the claimed function. The structure disclosed in the patents for performing the claimed function is

the microprocessor/CPU 31, the software stored in ROM 32, and controller 33. No other structure described in the patents performs the claimed function. The editing means does not correspond to a general-purpose microprocessor in a personal computer running software because the specification only discloses using a microprocessor in a special-purpose hardware unit, not in a personal computer. A ROM, like the one corresponding to this claim element, cannot be changed and therefore the editing program described in the patents cannot be changed. The specification never describes storing the program in a RAM, random access memory, or in other type of storage unit.

47. The claim language for the "recording means" does not identify either structure for performing the claimed function or a known class of structures for performing the claimed function. The patent identifies the internal elements of the audio/video recording unit (AVRU) 11 with media 23, which is a video recorder using magnetic tape or an optical disc, and mechanical shunt switch 48' as performing the claimed function. No other structure described in the patents performs the claimed function.

48. "Substantially shorter" as used in the claims of the '705 patent refers to magnitude, not approximation. One of skill in the art would understand that the examples Burst gave in the specification and file history of "substantially shorter" transmission time more than two orders of magnitude – *i.e.*, 1% –less than the normal view time. The specification explains "The fiber optic line carries digital signals in the form of light waves over great distances with a high degree of accuracy and reliability and at a high speed (e.g., about 200 megabytes/second)." '705 patent, Col. 7, Line 66 – Col. 8, Line 2. The patent also states that a typical feature length film could be compressed into as little as 250 megabytes. '705 patent, Col. 5, Lines 26-30 ("if no data compression technique is used, it would take approximately 51.03 gigabytes to store a 2

hour video program, but by using the above compression techniques, it is estimated that memory 13 will require only 250 megabytes"). 250 megabytes was an extraordinary amount of RAM memory in 1988. Even personal computer hard drives only had capacity measured in the tens of megabytes. Given such a combination, the transmission time of a feature length, 2 hour, film would be on the order of 1.25 seconds under these conditions. Even using a more realistic figure of about 5 gigabytes for modern DVD would result in a transmission time of significantly less than one minute– 25 seconds in this example. The specification of the '705 patent describes transmission examples in which videos are transmitted in a few seconds or minutes. Col. 11, Lines 44-49 ("Thus, the invention can be used to receive and transmit programs via microwaves at an accelerated rate similar to and at least as fast as, the transmission and reception of programs over optical fibers. This feature allows transmission and reception of programs in a few minutes or seconds using currently available technology.").

49. Consistent with the specification, the file history of the '995 patent describes a transmission time of 5 to 30 seconds for a two hour video. (March 20, 1990 Amendment A at 20) In this example, the compressed transmission time is 0.07-0.4% of the normal viewing time of the video.

50. Burst's expert, Dr. Stevenson, testified in his deposition that "substantially shorter" referred to short enough to influence system design by requiring a larger buffer. This construction is technologically vague because Dr. Stevenson makes several assumptions about the video system that are not required by any part of the patent claims or specification. It would be possible to use other components and techniques that would lead to different conclusions about what influences the system design and thus what would, as Dr. Stevenson reads this term, constitute a "substantially shorter" transmission time.

51. To one of ordinary skill in the art, the term "transceiver" refers to a receiver and transmitter that are combined together in a common housing. They use common circuits for both transmitting and receiving. The IEEE Authoritative Dictionary of IEEE Standards Terms, Seventh Edition (2000) defines "transceiver" as:

> transceiver( 1) (data transmission) the combination of radio transmitting and receiving equipment in a common housing, usually for portable or mobile use, and employing common circuit components for both transmitting and receiving. (PE) 599-1985w

My understanding as one of skill in the art is that the patents at issue use the term transceiver consistently with this definition.

52. As one of skill in the art, I understand "editing" to be using the "time compressed representation of said audio/video source information" as the original data and changing the sequence or appearance of that information or adding or subtracting information to create the "edited time compressed representation of said audio/video source information." Editing is not compiling a list or compilation of audio/video programs like a playlist or any other collection of information about the audio/video program. Editing refers instead to changing the content of the audio/video program.

53. A playlist identifies a sequence of files to be played back. A playlist does not contain the audio/video material itself but simply identifies the material to be played in a particular order. A playlist is not an edited version of the original material. In fact, there is no original material because the playlist simply identifies existing material to be played. The video material identified on the playlist never resides on a storage device as a whole. The playlist just causes software to go get files in a particular order.

54. It is certainly technically feasible to develop a system that processes video programs as a whole and sometimes it is desirable to do so. Not only is it possible to acquire all the video data before compressing it, some algorithms actually require input from all the data before commencing the compression process. One such situation is called vector quantization, or VQ. Such VQ algorithms can use customized codebooks that are specifically tailored for the video material being watched. Custom-codebook VQ algorithms scan through the input video data before selecting codebook entries optimized for that particular source material. Vector quantization benefits from custom-codebook generation optimized for the particular program to be viewed. It is necessary to preview the whole video to construct an optimal codebook.

55. I declare under penalty of perjury that the foregoing is true and correct. Executed on January 12, 2004.

Brian Von Herzen

Brian Von Herzen